On second order elliptic equations with a small parameter

Mark Freidlin* , Wenqing Hu†

Abstract

The Neumann problem with a small parameter

$$\left(\frac{1}{\varepsilon}L_0 + L_1\right) u^\varepsilon(x) = f(x) \text{ for } x \in G, \ \left.\frac{\partial u^\varepsilon(x)}{\partial \gamma^\varepsilon(x)}\right|_{\partial G} = 0$$

is considered in this paper. The operators L_0 and L_1 are self-adjoint second order operators. We assume that L_0 has a non-negative characteristic form and L_1 is strictly elliptic. The reflection is with respect to inward co-normal unit vector $\gamma^\varepsilon(x)$. The behavior of $\lim\limits_{\varepsilon\downarrow 0} u^\varepsilon(x)$ is effectively described via the solution of an ordinary differential equation on a tree. We calculate the differential operators inside the edges of this tree and the gluing condition at the root. Our approach is based on an analysis of the corresponding diffusion processes.

Keywords: second order equations with a small parameter, equations with non-negative characteristic form, diffusion processes on a graph, averaging principle.

2010 Mathematics Subject Classification Numbers: 35J57, 35J70, 60J60.

1 Introduction

Let G be a bounded domain in $\mathbb{R}^d$ with the smooth boundary ∂G,

$$L_k u(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}^{(k)}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i^{(k)}(x)\frac{\partial u}{\partial x_i} \ , \ k = 0, 1 \ , \ x \in \mathbb{R}^d \ .$$

The coefficients are assumed to be smooth enough, say, in $\mathbf{C}^{(2)}(\mathbb{R}^d)$, i.e., having continuous second derivatives.

Boundary problems for the operator $L_\varepsilon = L_0 + \varepsilon L_1$ in the domain G and corresponding initial-boundary problems for the equation $\dfrac{\partial u^\varepsilon(t,x)}{\partial t} = L_\varepsilon u^\varepsilon$, $t > 0$, $x \in G$, are considered. The operator L_ε is assumed to be elliptic for $\varepsilon > 0$. One can study the limiting behavior of solutions of stationary problems as $\varepsilon \downarrow 0$ and the limiting behavior of solutions of initial-boundary problems as $\varepsilon \downarrow 0$ and $t \to \infty$.

*Department of Mathematics, University of Maryland at College Park, mif@math.umd.edu

†Department of Mathematics, University of Maryland at College Park, huwenqing@math.umd.edu

If the operator L_0 is elliptic in $G \cup \partial G$, the problem is simple: u^ε converges to the solution of corresponding problem for the operator L_0. In the case of degenerate operator L_0, situation is more complicated, and the question was considered in numerous papers. First, the case of first order operator L_0 was considered: $L_0 = b^{(0)}(x) \cdot \nabla$, $b^{(0)}(x) = (b_1^{(0)}(x), ..., b_d^{(0)}(x))$. N.Levinson [14] showed in 1950-th that, if the characteristics of L_0 (e.g., trajectories of the dynamical system $\dot{X}_t = b^{(0)}(X_t)$ in $\mathbb{R}^d$) leave the domain G in finite time and cross the boundary in a regular way, then the solution of the Dirichlet problem $L_\varepsilon u^\varepsilon = 0$, $x \in G$, $u^\varepsilon(x)|_{\partial G} = \psi(x)$, converges as $\varepsilon \downarrow 0$ to the solution of degenerate equation $L_0 u^0(x) = 0$, $x \in G$, with the boundary condition $\psi(x)$ ($\psi(x)$ is assumed to be continuous) on the part of ∂G through which the characteristics leave the domain. Such a solution $u^0(x)$ is unique.

Most of subsequent results concerning this problem were obtained by probabilistic methods. With each operator L_ε, $\varepsilon \geq 0$, one can (see [3], notice that the coefficients of $a_{ij}^{(k)}(x)$ are in $\mathbf{C}^{(2)}(\mathbb{R}^d)$) connect a diffusion process $\widetilde{X}_t^\varepsilon$ in $\mathbb{R}^d$ defined by the stochastic differential equation

$$
\begin{aligned}
&\dot{\widetilde{X}}_t^\varepsilon = b^{(0)}(\widetilde{X}_t^\varepsilon) + \varepsilon b^{(1)}(\widetilde{X}_t^\varepsilon) + \sigma^{(0)}(\widetilde{X}_t^\varepsilon)\dot{W}_t^0 + \sqrt{\varepsilon}\sigma^{(1)}(\widetilde{X}_t^\varepsilon)\dot{W}_t^1 \ , \\
&\widetilde{X}_0^\varepsilon = x \in \mathbb{R}^d \ , \ t > 0 \ , \ \sigma^{(k)}(x)(\sigma^{(k)}(x))^* = (a_{ij}^{(k)}(x)) = a^{(k)}(x) \ , \ k = 0, 1 \ .
\end{aligned}
$$

Here W_t^0 and W_t^1 are independent Wiener processes in $\mathbb{R}^d$. Then the solution of the Dirichlet problem for the equation $L_\varepsilon u^\varepsilon(x) = 0$, $x \in G$, and of the initial boundary problem for $\dfrac{\partial u^\varepsilon(t,x)}{\partial t} = L_\varepsilon u^\varepsilon(t,x)$ can be represented as expectations of corresponding functionals of $\widetilde{X}_t^\varepsilon$. The trajectories $\widetilde{X}_t^\varepsilon$, in a sense, play the same role as characteristics in the case of first order operator L_0. Using these representations and studying limiting behavior of process $\widetilde{X}_t^\varepsilon$ one can describe the limiting behavior of the boundary problems (see [4], [6]).

If problems with the Neumann boundary conditions are considered, one can use the corresponding diffusion process with reflection on the boundary (see, for instance, [4, §2.5]). Various cases of first order operators L_0 not satisfying Levinson's conditions were examined using the probabilistic approach (see [4], [6] and the references therein).

If the operator L_0 has terms with second derivatives, one can introduce a generalized Levinson condition ([4, §4.2]). Under this condition the equation $L_0 u^0(x) = 0$, $x \in G$, with appropriate Dirichlet type boundary conditions has a unique solution, and the solution $u^\varepsilon(x)$ of the Dirichlet problem for equation $L_\varepsilon u^\varepsilon(x) = 0$, $x \in G$ converges to $u^0(x)$ as $\varepsilon \downarrow 0$. The difference with the classical Levinson case is just in the rate of convergence: under mild additional assumptions $|u^\varepsilon(x) - u^0(x)| < \varepsilon^\gamma$ for some $\gamma > 0$ and $0 < \varepsilon << 1$, but for any $\gamma' > 0$ one can find L_0 with infinitely differentiable coefficients non-degenerating on ∂G such that $|u^\varepsilon(x) - u^0(x)|$ is greater than $\varepsilon^{\gamma'}$ at a point $x \in G$ and $0 < \varepsilon << 1$.

A convenient way to specify the degeneration of L_0 is given by the conservation laws. A function $H(x)$ is called a first integral for the process X_t^0 corresponding to L_0 if $\mathbf{P}_x(X_t^0 \in S(H(x))) = 1$ for all $t \geq 0$ and $x \in \mathbb{R}^d$, where $S(z) = \{y \in \mathbb{R}^d : H(y) = z\}$; here and below the subscript $x \in \mathbb{R}^d$ in the probability $\mathbf{P}_x$ or expected value $\mathbf{E}_x$ means that the trajectory of the process starts at the point x.

We consider in this paper self-adjoint operators L_0 and L_1:

$$L_k u(x) = \frac{1}{2}\nabla \cdot (a^{(k)}(x)\nabla u(x)) \ .$$

Then a smooth function $H(x)$ is a first integral for the process $\widetilde{X}_t^0$ (for the corresponding operator L_0) if and only if $a^{(0)}(x)\nabla H(x) \equiv 0$. In general, the process $\widetilde{X}_t^0$ can have several independent smooth first integrals. To restrict ourselves to the case of one smooth first integral we assume that $\mathbf{e}\cdot(a^{(0)}(x)\mathbf{e}) \geq \underline{a}(x)|\mathbf{e}|^2_{\mathbb{R}^d}$ for each $\mathbf{e} \in \mathbb{R}^d$ such that $\mathbf{e}\cdot\nabla H(x) = 0$: It is assumed that $\underline{a}(x)$ is smooth and strictly positive if x is not a critical point of $H(x)$; if x_0 is a critical point, $a^{(0)}(x_0) = 0$ and $\underline{a}(x_0) = 0$.

To be specific we consider the Neumann problem

$$\frac{1}{\varepsilon}L_\varepsilon u^\varepsilon = \left(\frac{1}{\varepsilon}L_0 + L_1\right) u^\varepsilon(x) = f(x) \ , \quad \frac{\partial u^\varepsilon(x)}{\partial \gamma^\varepsilon(x)}\bigg|_{\partial G} = 0 \ ; \tag{1.1}$$

$\gamma^\varepsilon(x)$ here is the inward co-normal unit vector to ∂G corresponding to L_ε. Let X_t^ε be the process in $G \cup \partial G$ governed by the operator inside G with reflection along the co-normal to ∂G. Since L_ε is self-adjoint, the Lebesgue measure is invariant for the process X_t^ε, and the problem (1.1) is solvable for continuous $f(x)$ such that $\int_G f(x)dx = 0$. Together with the last condition, we assume that L_1 is not degenerate in $G \cup \partial G$, so that to single out a unique solution of (1.1), we shall fix the value of $u^\varepsilon(x)$ at a point $x_O \in G \cup \partial G$ which is fixed the same for all $\varepsilon > 0$. We let $u^\varepsilon(x_O) = 0$.

Then the solution of problem (1.1) can be written in the form (see, for instance [4])

$$u^\varepsilon(x) = -\int_0^\infty \mathbf{E}_x f(X_t^\varepsilon)dt + \int_0^\infty \mathbf{E}_{x_O} f(X_t^\varepsilon)dt \ . \tag{1.2}$$

If the first integral $H(x)$ has in $G \cup \partial G$ no critical points, one can describe the $\lim_{\varepsilon\downarrow 0} u^\varepsilon(x)$ in the way similar to [9]: One shall introduce a graph $\mathbb{G}$ corresponding to the set of connected components of the intersections of the level sets of $H(x)$ within G. A boundary problem on $\mathbb{G}$ with appropriate gluing conditions at the vertices can be formulated, and the solution of this problem defines $\lim_{\varepsilon\downarrow 0} u^\varepsilon(x)$. If the function $H(x)$ has saddle points inside G, additional branchings in the graph appear. The gluing conditions at these new vertices can be calculated using the results of [5].

All mentioned above results concern the case when the rank of $a^{(0)}(x)$ is constant and equal to $d-1$ for all $x \in G \cup \partial G$ except the critical points of $H(x)$. In this paper, we consider the case when L_0 is non-degenerate in a connected subdomain $\mathcal{E} \subset G$, and we let

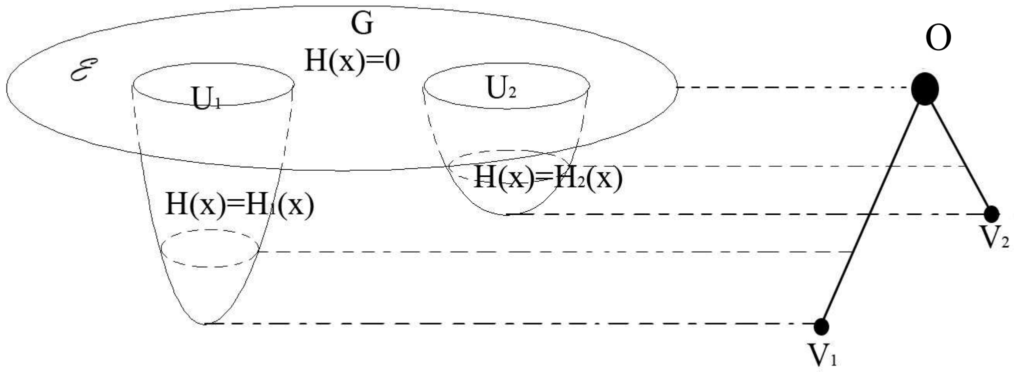

Fig. 1.

$H(x)$ be equal to a constant on $\mathcal{E}$. Outside $\mathcal{E}$ the first integral $H(x)$ has a finite number of critical points (see Fig.1). For convenience of presentation, we shall then introduce several first integrals $H_k(x)$ $(k = 1, ..., r)$ for each of the connected components $U_1, ..., U_r$ on which L_0 is degenerate. We shall let $H(x) = H_k(x)$ for $x \in U_k$. A more concrete setup of the problem is in Section 2. Existence of the domain $\mathcal{E}$ where the operator L_0 is not degenerate leads to more general gluing conditions. The limiting process on the graph spends a positive time at the vertex corresponding to $\mathcal{E}$.

Let $S(z) = \{x \in G \cup \partial G : H(x) = z\}$. The graph $\mathbb{G}$ is the result of identification of points of each connected component of every level set $S(z)$. Let $\mathfrak{Y} : G \cup \partial G \to \mathbb{G}$ be the identification mapping. We call $\mathfrak{Y}(x)$ the projection of x onto $\mathbb{G}$. We consider the projection $Y_t^\varepsilon = \mathfrak{Y}(X_t^\varepsilon)$ of the process X_t^ε on $\mathbb{G}$ and prove that processes Y_t^ε on $\mathbb{G}$ converge weakly in the space of continuous functions $[0, T] \to \mathbb{G}$ to a diffusion process Y_t on $\mathbb{G}$. The process Y_t is defined by a family of differential operators, one on each edge, and by gluing conditions at the vertices. We calculate the operators and the gluing condition. The function $u^0(x) = \lim\limits_{\varepsilon \downarrow 0} u^\varepsilon(x)$ is constant on each connected component of every level set of $H(x)$: $u^0(x) = v(\mathfrak{Y}(x))$. We formulate a boundary problem for the function $v(y)$, $y \in \mathbb{G}$, which has a unique solution, and actually can be solved explicitly.

The organization of this paper is as follows: Section 2 sets up the problem and gives the main results. Section 3 is devoted to the proof of the main results in Section 2. Section 4 proves auxiliary results needed in Section 3.

2 Main results

Let us first speak about our assumptions.

Suppose we have a bounded domain $G \subset \mathbb{R}^d$, with smooth boundary ∂G. Let L_0

be a self-adjoint operator

$$L_0 u(x) = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}^{(0)}(x)\frac{\partial u(x)}{\partial x_j}\right) = \frac{1}{2}\nabla\cdot(a^{(0)}(x)\nabla u(x)) \ .$$

Let $U_1, ..., U_r$ be several regions inside G. They are simply connected open sets and they do not intersect each other. Let us assume, that the matrix $a^{(0)}(x) \equiv (a_{ij}^{(0)}(x))_{1\le i,j\le d}$ is strictly elliptic on $\mathcal{E} = [G]\backslash(\cup_{k=1}^{r}[U_k])$ (here $[D]$ is the closure of a domain D). For $x \in [\mathcal{E}]$, the coefficients $a_{ij}^{(0)}(x), 1 \le i,j \le d$ are assumed to be in $\mathbf{C}^{(3)}([\mathcal{E}])$.

Let us discuss the case when $x \in \cup_{k=1}^{r}[U_k]$. For each $k = 1, ..., r$ and $x \in [U_k]$, the coefficients $a_{ij}^{(0)}(x), 1 \le i,j \le d$ are assumed to be in $\mathbf{C}^{(3)}([U_k])$. We assume that the matrix $(a_{ij}^{(0)}(x))_{1\le i,j\le d}$ is degenerate on $\cup_{k=1}^{r}[U_k]$. To specify this degeneration, we assume that within each $[U_k]$ there is a function H_k which is a first integral of the (degenerate) operator L_0, i.e., $a^{(0)}(x)\nabla H_k(x) = 0$ for $x \in [U_k]$. Let H_k have only one minimum m_k inside U_k. (We can always make this assumption since if m_k is a maximum we work with $-H_k$ instead of H_k.) Let $x_k(m_k)$ be the point in U_k corresponding to the minimum m_k. This minimum is assumed to be non-degenerate, i.e., the matrix $\left(\frac{\partial^2 H_k}{\partial x_i \partial x_j}(x_k(m_k))\right)_{1\le i,j\le d}$ is positive definite. Since the choice of H_k is up to a constant we can assume that $H_k = 0$ on ∂U_k. For $h \in (m_k, 0]$ the level surfaces $C_k(h) = \{x \in U_k : H_k(x) = h\}$ of the functions H_k inside U_k are closed surfaces of dimension $(d-1)$ and the operator L_0 is non-degenerate on $C_k(h)$. Let $\gamma_k = \partial U_k = C_k(0)$. A non-degeneracy condition of $a^{(0)}(x)$ on $C_k(h)$ is assumed: for any vector $\mathbf{e} \in \mathbb{R}^d$ such that $\mathbf{e}\cdot\nabla H_k = 0$ we have $\mathbf{e}\cdot(a^{(0)}(x)\mathbf{e}) \ge \underline{a}(x)|\mathbf{e}|^2_{\mathbb{R}^d}$ for some $\underline{a}(x) > 0$ and $x \ne x_k(m_k)$. We set $a^{(0)}(x_k(m_k)) = 0$ and $C_k(m_k) = \{x_k(m_k)\}$. We assume that the level surfaces $C_k(h)$ for $h \in (m_k, 0]$ divide $U_k\backslash\{x_k(m_k)\}$ into pieces of $(d-1)$-dimensional surfaces.

For simplicity of presentation we will assume that $\nabla H_k(x) \ne 0$ for $x \in \gamma_k$. One can introduce a global first integral $H(x)$ on $[G]$ as in Section 1: $H(x) = H_k(x)$ for $x \in U_k$ and $H(x) = 0$ for $x \in [\mathcal{E}]$. However, this function $H(x)$ is not smooth at $\cup_{k=1}^{r}\gamma_k$ but this is only a result of non-essential technical assumptions.

Let $\gamma = \cup_{k=1}^{r}\gamma_k$. We assume that the order of degeneracy is given by the condition that for a certain unit vector field $\mathbf{e}_d(x)$ in a small neighborhood of $\cup_{k=1}^{r}[U_k]$ we have

$$\mathrm{const}_1\cdot\mathrm{dist}^2(x,\gamma) \le \mathbf{e}_d(x)\cdot(a^{(0)}(x)\mathbf{e}_d(x)) \le \mathrm{const}_2\cdot\mathrm{dist}^2(x,\gamma)$$

for some const_1, $\mathrm{const}_2 > 0$. The distance $\mathrm{dist}(x,\gamma)$ is the Euclidean distance between x and γ. The vector field $\mathbf{e}_d(x) = \dfrac{\nabla H_k}{|\nabla H_k|_{\mathbb{R}^d}}$ for $x \in \gamma_k$.

In particular, our assumptions imply that the matrix $a^{(0)}(x)$ has rank d in $\mathcal{E}$ and rank $(d-1)$ in $\cup_{k=1}^{r}[U_k]$. However, the coefficients $a_{ij}^{(0)}(x)$, $1 \le i,j \le d$ are only in $\mathbf{C}^{(1)}$

for $x \in \gamma$. We notice that in this case results of [3] do not apply. We shall then assume that there is a decomposition $a^{(0)}(x) = \sigma^{(0)}(x)(\sigma^{(0)}(x))^*$ for all $x \in [G]$. The square matrix $\sigma^{(0)}(x)$ has bounded Lipschitz continuous terms.

We shall assume, that the operator L_1 governing the perturbation is self-adjoint and strictly elliptic within $[G]$:

$$L_1 u(x) = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left(a_{ij}^{(1)}(x) \frac{\partial u(x)}{\partial x_j} \right) = \frac{1}{2} \nabla \cdot (a^{(1)}(x) \nabla u(x)) \ .$$

Again we denote the matrix $a^{(1)}(x) \equiv (a_{ij}^{(1)}(x))_{1 \le i,j \le d}$ and we assume that the terms $a_{ij}^{(1)}(x)$ are in class $\mathbf{C}^{(2)}(\mathbb{R}^d)$. In this case results of [3] apply and we have $a^{(1)}(x) = \sigma^{(1)}(x)(\sigma^{(1)}(x))^*$ for all $x \in [G]$. The square matrix $\sigma^{(1)}(x)$ have bounded Lipschitz continuous terms.

Let us put a Neumann boundary condition with respect to co-normal unit vector $\gamma^\varepsilon(x)$ pointing inward on ∂G. Let X_t^ε be the diffusion process in $[G]$, corresponding to the operator $\dfrac{1}{\varepsilon} L_0 + L_1$ inside G with co-normal reflection at ∂G. We see that Lebesgue measure is invariant for the process X_t^ε.

Let us then speak about the results.

We construct a graph $\mathbb{G}$ as follows. The graph $\mathbb{G}$ has r edges $I_1, ..., I_r$ joined together at one vertex O. Let the other endpoint of I_k be V_k. Let us write $I_k = [m_k, 0]$. The coordinate (k, H_k) is a global coordinate on $\mathbb{G}$. The root O corresponds to all $(k, 0)$ for $k = 1, ..., r$. Let us introduce an identification map $\mathfrak{Y} : [G] \to \mathbb{G}$: for $x \in [\mathcal{E}]$ we have $\mathfrak{Y}(x) = O$ and for $x \in U_k$ we have $\mathfrak{Y}(x) = (k, H_k(x))$. Let the process $Y_t^\varepsilon = \mathfrak{Y}(X_t^\varepsilon)$. We are going to prove, that as $\varepsilon \downarrow 0$ the processes Y_t^ε converge weakly in the space $\mathbf{C}_{[0,T]}(\mathbb{G})$ to a Markov process Y_t on $\mathbb{G}$.

The process Y_t is defined as follows. It is a diffusion process on the graph $\mathbb{G}$ with generator A and the domain of definition $D(A)$. Inside each I_k it is governed by an operator $\mathcal{L}_k$ defined as

$$\mathcal{L}_k f(k, H_k) = \frac{1}{2} M_k^{-1}(H_k) \frac{d}{dH_k} \left(M_k(H_k) \overline{a^{(1)}}(H_k) \frac{df}{dH_k} \right) \ .$$

Here

$$\overline{a^{(1)}}(h) = M_k^{-1}(h) \int\limits_{C_k(h)} \frac{(a^{(1)}(x) \nabla H_k(x), \nabla H_k(x))}{|\nabla H_k(x)|} d\sigma \ ,$$

and normalizing factor

$$M_k(h) = \int\limits_{C_k(h)} \frac{d\sigma}{|\nabla H_k(x)|} \ .$$

The notation $d\sigma$ denotes the integral with respect to the area element on $C_k(h)$.

We set $Af = \mathcal{L}_k f$ for $(k, H_k) \in (I_k)$ ((I_k) is the interior of the interval I_k). Let the limit $\lim\limits_{(k,H_k)\to O} Af(k, H_k)$ be finite and independent of k. This limit is set to be $Af(O)$. The domain of definition $D(A)$ of the operator A consists of those functions f that are twice continuously differentiable inside each I_k having the limit $\lim\limits_{H_k\to 0} \dfrac{\partial f}{\partial H_k}(k, H_k)$. These functions satisfy the gluing condition at the vertex O:

$$0 = \text{Volume}(\mathcal{E}) \cdot Af(O) + \frac{1}{2}\sum_{k=1}^{r} p_k \cdot \lim_{H_k\to 0} \frac{\partial f}{\partial H_k}(k, H_k) \ . \tag{2.1}$$

Here Volume($\mathcal{E}$) is the d-dimensional volume of the domain $\mathcal{E}$ and

$$p_k = \int_{\gamma_k} \frac{(a^{(1)}(x)\nabla H_k(x), \nabla H_k(x))}{|\nabla H_k(x)|} d\sigma \ .$$

For the exterior vertices $V_1, ..., V_r$ no additional assumptions are to be imposed on the behavior of the function f in the domain $D(A)$.

It was proved in [7] the the process Y_t exists and is a strong Markov process on the graph $\mathbb{G}$.

We have

Theorem 2.1. *As $\varepsilon \downarrow 0$ the processes Y_t^ε converge weakly to Y_t in $\mathbf{C}_{[0,T]}(\mathbb{G})$.*

Let μ_x^ε be the distribution of the trajectory $Y_t^\varepsilon = \mathfrak{Y}(X_t^\varepsilon)$ starting from a point $x \in [G]$ in the space $\mathbf{C}_{[0,T]}(\mathbb{G})$: for each Borel subset $B \subseteq \mathbf{C}_{[0,T]}(\mathbb{G})$ we set $\mu_x^\varepsilon(B) = \mathbf{P}_x(Y_\bullet^\varepsilon \in B)$. Similarly, for each $y \in \mathbb{G}$ we let μ_y^0 be the distribution of Y_t in the space $\mathbf{C}_{[0,T]}(\mathbb{G})$ with $\mu_y^0(B) = \mathbf{P}_y(Y_\bullet \in B)$. Theorem 2.1 can be reformulated as

Theorem 2.2. *For every $x \in [G]$ and every $T > 0$ the distribution μ_x^ε converges weakly to $\mu_{\mathfrak{Y}(x)}^0$ as $\varepsilon \downarrow 0$. For every bounded continuous functional F on $\mathbf{C}_{[0,T]}(\mathbb{G})$ we have*

$$\mathbf{E}_x F(Y_\bullet^\varepsilon) \to \mathbf{E}_{\mathfrak{Y}(x)} F(Y_\bullet)$$

as $\varepsilon \downarrow 0$.

The process Y_t^ε can be viewed as the slow component of the process X_t^ε. The fast component Z_t^ε of X_t^ε is a process governed by the operator $\dfrac{1}{\varepsilon}L_0$. The process Z_t^ε moves on $\mathfrak{Y}^{-1}(y)$ for each $y \in \mathbb{G}$: it is moving on $[\mathcal{E}]$ when $y = O$ and it is moving on $C_k(H_k)$ when $y = (k, H_k)$. Since Lebesgue measure is invariant for the process X_t^ε, the fast component Z_t^ε, as $\varepsilon > 0$ is small, has, approximately, a distribution with density $\dfrac{1}{\text{Volume}(\mathcal{E})}$ on $[\mathcal{E}]$ (with respect to Lebesgue measure on $\mathbb{R}^d$) and $\dfrac{1}{M_k(H_k)}\dfrac{1}{|\nabla H_k|_{\mathbb{R}^d}}$ on

$C_k(H_k)$ (with respect to the area element $d\sigma$ on $C_k(H_k)$). Using this we can formulate the above two theorems in terms of differential equations:

Theorem 2.3. *Consider the Neumann problem*

$$\frac{1}{\varepsilon}L_\varepsilon u^\varepsilon(x) = \left(\frac{1}{\varepsilon}L_0 + L_1\right) u^\varepsilon(x) = f(x) \text{ for } x \in G \ , \ \left.\frac{\partial u^\varepsilon(x)}{\partial \gamma^\varepsilon(x)}\right|_{x\in\partial G} = 0$$

with a Hölder continuous function $f(x)$ satisfying $\int_G f(x)dx = 0$. Let $u^\varepsilon(x_O) = 0$ for some $x_O \in G$. Then we have

$$\lim_{\varepsilon\downarrow 0} u^\varepsilon(x) = v(\mathfrak{Y}(x))$$

where $v(y)$ is a continuous function on $\mathbb{G}$ such that

$$\mathcal{L}_k v(y) = -\bar{f}(y) \text{ for } y \in (I_k) \ , \ k = 1, ..., r \ .$$

Here

$$\bar{f}(y) = \frac{1}{\text{Volume}(\mathcal{E})}\int_{\mathcal{E}} f(x)dx$$

when $y = O$ and

$$\bar{f}(y) = \frac{1}{M_k(H_k)}\int_{C_k(H_k)} f(x)\frac{d\sigma}{|\nabla H_k(x)|_{\mathbb{R}^d}}$$

when $y = (k, H_k)$. The function $v(y)$ satisfies the gluing condition (2.1) and $v(\mathfrak{Y}(x_O)) = 0$. Such a function $v(y)$ is unique.

3 Proof of Theorem 2.1

The **Proof** of Theorem 2.1 follows the arguments of [6], [8], [1] and [2].

Heuristically, the idea of [2] can be explained as follows. The process X_t^ε moves within $[G]$ and has Lebesgue measure as its invariant measure. Since the process X_t^ε has a "fast" component governed by the operator $\frac{1}{\varepsilon}L_0$, it will spend a positive amount of time proportional to Volume($\mathcal{E}$) within $\mathcal{E}$ as $\varepsilon \downarrow 0$. As we project X_t^ε onto the graph $\mathbb{G}$ and the whole ergodic component $\mathcal{E}$ corresponds to O, the limiting process Y_t has a boundary condition with a "delay" at O. (We recommend a nice article [12] and a brief summary [13, §5.7] about this boundary condition.) Our gluing condition (2.1) ensures that the process Y_t has an invariant measure on $\mathbb{G}$ that agrees with the Lebesgure measure on $[G]$. We also refer to [6, Ch.8, pp. 347–350] for an explanation of this.

Let us first introduce some notations. Below we will often suppress the small parameter ε and it could be understood directly from the context. Let $\overline{\gamma}_k = C_k(-\varepsilon^{1/2})$ and $\overline{\gamma} = \cup_{k=1}^r \overline{\gamma}_k$. Let σ be the first time when the process X_t^ε hits γ. Let τ be the first time when the process X_t^ε hits $\overline{\gamma}$. Let $\sigma_0 = \sigma$. Let τ_n be the first time following σ_n when

the process reaches $\overline{\gamma}$. For $n \geq 1$ let σ_n be the first time after τ_{n-1} when the process X_t^ε reaches γ.

Let $\sigma^* \in \{\sigma_0, \sigma_1, ...\}$ and we denote by $m_{\sigma^*}^x$ the measure on γ induced by $X_{\sigma^*}^\varepsilon$ starting at x. That is,

$$m_{\sigma^*}^x(A) = \mathbf{P}_x(X_{\sigma^*}^\varepsilon \in A) \ , \ A \in \mathcal{B}(\gamma) \ .$$

Let $\nu(\bullet)$ be the invariant measure of the induced chain $X_{\sigma_n}^\varepsilon$ on γ. The key lemma of [2] is the following

Lemma 3.1. *Let $x \in [\mathcal{E}]$. For each $\delta > 0$ and all sufficiently small ε there is a stopping time σ^* which may depend on δ, ε and x and such that*

$$\mathbf{E}_x \sigma^* \leq \delta \ , \tag{3.1}$$

$$\sup_{x \in \gamma} Var(m_{\sigma^*}^x(dy) - \nu(dy)) \leq \delta \ , \tag{3.2}$$

where Var is the total variation of the signed measure.

Our proof of this lemma is a bit simpler than that of [2].

Proof. We will prove, in Lemma 4.11 that $X_{\sigma_n}^\varepsilon$ satisfies the Doeblin condition on γ uniformly in ε. This implies that one can choose an N depending on δ but independent of ε such that the distribution of $X_{\sigma_N}^\varepsilon$ is δ-close to the invariant measure $\nu(\bullet)$ on γ. That is, as we set $\sigma^* = \sigma_N$ the condition (3.2) is satisfied.

We are going to prove in Lemmas 4.8, 4.9, 4.10, respectively, that

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in [\mathcal{E}]} \mathbf{E}_x \sigma = 0 \ , \tag{3.3}$$

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \gamma} \mathbf{E}_x \tau = 0 \ , \tag{3.4}$$

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \overline{\gamma}} \mathbf{E}_x \sigma = 0 \ , \tag{3.5}$$

uniformly in ε. We can write $\sigma_N = \sum_{i=1}^{N} [(\sigma_i - \tau_{i-1}) + (\tau_{i-1} - \sigma_{i-1})] + \sigma_0$. For each $i = 1, ..., N$ the random variable $\sigma_i - \tau_{i-1}$ has the same distribution as σ for the process X_t^ε starting at some point on $\overline{\gamma}$; similarly, the random variable $\tau_{i-1} - \sigma_{i-1}$ has the same distribution as τ for the process X_t^ε starting at some point on γ. The results (3.3), (3.4), (3.5) imply that as ε is small (notice that N is fixed at this stage) we can choose $\sigma^* = \sigma_N$ and the condition (3.1) is also satisfied. $\square$

Proof of Theorem 2.1. The proof of Theorem 2.1 is the same as the proof of Lemma 2.1 (including the proof of Lemma 3.4) stated in [2] using the above Lemma 3.1. For the sake of completeness let us briefly repeat it here. Reasoning as in [1], [2], [6], [8], it suffices to prove that for a function $f \in D(A)$, for every $T > 0$ and uniformly in $x \in [G]$ we have

$$\mathbf{E}_x \left[f(H(X_T^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^T Af(H(X_s^\varepsilon))ds \right] \to 0$$

as $\varepsilon \downarrow 0$. Here $H(x) = H_k(x)$ if $x \in U_k$ and $H(x) = 0$ if $x \in [\mathcal{E}]$. Let us replace the time interval $[0, T]$ by a larger one $[0, \widetilde{\sigma}]$, where $\widetilde{\sigma}$ is the first of the stopping times σ_n which is greater than or equal to T: $\widetilde{\sigma} = \min\limits_{n:\sigma_n > T} \sigma_n$. Let $\widetilde{\sigma} = \sigma_{\widetilde{n}+1}$. We have

$$\begin{aligned}
&\mathbf{E}_x \left[f(H(X_T^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^T Af(H(X_s^\varepsilon))ds \right] \\
&= \mathbf{E}_x \left[f(H(X_\sigma^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^\sigma Af(H(X_s^\varepsilon))ds \right] + \mathbf{E}_x \sum_{k=0}^{\widetilde{n}} \int_{\sigma_k}^{\sigma_{k+1}} Af(H(X_s^\varepsilon))ds - \\
&\quad - \mathbf{E}_x \mathbf{E}_{X_T^\varepsilon} \left[f(H(X_\sigma^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^\sigma Af(H(X_s^\varepsilon))ds \right] \\
&= (I) + (II) - (III) \ .
\end{aligned}$$

If $x \in \cup_{k=1}^r U_k$ we have $|(I)| \to 0$ uniformly in x as $\varepsilon \downarrow 0$ due to averaging principle. If $x \in [\mathcal{E}]$ then $|(I)| \to 0$ uniformly in x as $\varepsilon \downarrow 0$ due to Lemma 4.8. In a similar way we see that $|(III)| \to 0$ uniformly in $x \in [G]$ as $\varepsilon \downarrow 0$.

Let $\alpha_k = \displaystyle\int_{\sigma_k}^{\sigma_{k+1}} Af(H(X_s^\varepsilon))ds$ and let $\beta_k = \displaystyle\sum_{n=0}^{\infty} \mathbf{E}_x(\alpha_{k+n}|\mathcal{F}_k)$ ($\mathcal{F}_k$ is the filtration generated by the process X_t^ε for $t \le \sigma_k$). We have $\mathbf{E}_x(\alpha_k - \beta_k + \beta_{k+1}|\mathcal{F}_k) = 0$ and therefore $\left(\displaystyle\sum_{k=1}^{n} (\alpha_k - \beta_k + \beta_{k+1}), \mathcal{F}_{n+1} \right)$, $n \ge 0$ is a martingale. Using the optimal sampling theorem we have

$$\mathbf{E}_x \sum_{k=0}^{\widetilde{n}} \alpha_k = \mathbf{E}_x \sum_{k=0}^{\widetilde{n}} (\alpha_k - \beta_k + \beta_{k+1}) + \mathbf{E}_x(\beta_0 - \beta_{\widetilde{n}+1}) = \mathbf{E}_x(\beta_0 - \beta_{\widetilde{n}+1}) \ .$$

The above argument shows that for the proof of $|(II)| \to 0$ uniformly in $x \in [G]$ as $\varepsilon \downarrow 0$ it suffices to prove $\sup\limits_{x\in\gamma} \left| \displaystyle\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n \right| \to 0$ uniformly in $x \in \gamma$ as $\varepsilon \downarrow 0$.

Let us first show that $\mathbf{E}_\nu \alpha_0 = 0$. By Lemma 4.11 the Markov chain $X_{\sigma_n}^\varepsilon$, $n \ge 0$ on γ is ergodic and has invariant measure ν. Therefore we have $\lim\limits_{n\to\infty} \dfrac{\sigma_n}{n} = \mathbf{E}_\nu \sigma_1$. By ergodicity of the process X_t^ε and self-adjointness of L_0 and L_1 we also have

$$\lim_{t\to\infty} \mathbf{E}_\nu \frac{1}{t} \int_0^t Af(H(X_s^\varepsilon))ds = \int_{[G]} Af(H(x))dx \ .$$

These two equalities imply that $\mathbf{E}_\nu \alpha_0 = \mathbf{E}_\nu \int_0^{\sigma_1} Af(H(X_s^\varepsilon))ds = (\mathbf{E}_\nu \sigma_1) \cdot \int_{[G]} Af(H(x))dx$.

We have

$$\int_{[G]} Af(H(x))dx = \text{Volume}(\mathcal{E}) \cdot Af(O) + \sum_{k=1}^{r} \int_{I_k} M_k(H_k)\mathcal{L}_k f(H_k)dH_k \ .$$

(The notations agree with those in the definition of the process Y_t.)

Since

$$\int_{I_k} M_k(H_k)\mathcal{L}_k f(H_k)dH_k = \int_{I_k} \frac{d}{dH_k}\left(M_k(H_k)\overline{a^{(1)}}(H_k)\frac{df}{dH_k}\right) dH_k = \frac{1}{2}p_k \cdot \lim_{H_k \to 0} \frac{df}{dH_k}(k, H_k) \ ,$$

we can use our boundary condition (2.1) to have $\int_{[G]} Af(H(x))dx = 0$ and therefore $\mathbf{E}_\nu \alpha_0 = 0$.

From the fact that $\mathbf{E}_\nu \alpha_0 = 0$ one first derives that $\sup\limits_{x\in\gamma} \mathbf{E}_x \alpha_n$ decays to 0 exponentially fast and therefore $\sup\limits_{x\in\gamma} \left|\sum\limits_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| < \infty$. It also gives, for $x \in \gamma$, that, for $\sigma^* \in \{\sigma_1, \sigma_2, ...\}$ we have

$$\left|\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \leq \|Af\|_\infty \cdot \mathbf{E}_x \sigma^* + \text{Var}(m_{\sigma^*}^x - \nu) \cdot \sup_{x\in\gamma} \left|\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \ .$$

Using Lemma 3.1 we see that for any $\delta > 0$ we can choose σ^* such that

$$\sup_{x\in\gamma} \left|\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \leq \|Af\|_\infty \cdot \delta + \delta \cdot \sup_{x\in\gamma} \left|\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \ ,$$

which proves that $\sup\limits_{x\in\gamma} \left|\sum\limits_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \to 0$ uniformly in $x \in \gamma$ as $\varepsilon \downarrow 0$. This implies that $|(II)| \to 0$ uniformly in $x \in [G]$ as $\varepsilon \downarrow 0$ and Theorem 2.1 follows. $\square$

4 Auxiliary results needed in the proof of Theorem 2.1

We establish in this section all the auxiliary results needed in Section 3 for the proof of Theorem 2.1.

Let us make some further geometric constructions. Since we assumed that all these U_k's do not intersect each other we see that for sufficiently small neighborhoods of these U_k's they also do not intersect each other. Without loss of generality let us speak about one of these U_k's. We remind that the matrix $a^{(0)}(x) = (a_{ij}^{(0)}(x))_{1\leq i,j\leq d}$ is non-negative definite inside $[G]$ and has rank d on $[G] \backslash \cup_{k=1}^r [U_k]$ and rank $(d-1)$ on $\cup_{k=1}^n [U_k]$. Since the operator L_0 is non-degenerate on $C_k(h)$ for $h \in (m_k, 0]$ we see that $a^{(0)}(x)\nabla H_k = 0$

on $C_k(h)$ and $\mathbf{e}\cdot(a^{(0)}(x)\mathbf{e}) \geq \underline{a}(x)|\mathbf{e}|^2_{\mathbb{R}^d}$ for any unit vector $\mathbf{e}$ tangent to $C_k(h)$. Here $\underline{a}(x) > 0$ for $x \in C_k(h)$ and $h \in (m_k, 0]$. The eigenvalue 0 for $a^{(0)}(x)$, $x \in C_k(0) = \gamma_k$ is simple and is the smallest in the spectrum of $a^{(0)}(x)$. By a transversality argument one can see that this eigenvalue will remain simple and is still the smallest one in the spectrum of all the matrices $a^{(0)}(x)$ as x belongs to a small neighborhood of U_k. As a consequence, the unit eigenvector $\mathbf{e}_d(x)$ (for different k it is different vector fields but for simplicity of notation we ignore that k in our notation) corresponding to this smallest eigenvalue is a $\mathbf{C}^{(3)}$ vector field in a neighborhood of U_k, with $\mathbf{e}_d(x) = \dfrac{\nabla H_k(x)}{|\nabla H_k(x)|_{\mathbb{R}^d}}$ for $x \in \gamma_k$. Let $X^x(t)$ be the integral curve of this vector field. We let $\dfrac{dX^x(t)}{dt} = \mathbf{e}_d(X^x(t))$, $X^x(0) = x \in \gamma_k$. As we are working within a small neighborhood of γ_k and $\mathbf{e}_d(x)$ in this neighborhood is a $\mathbf{C}^{(3)}$ vector field, being transversal to γ_k when $x \in \gamma_k$, we see that for $t \in [0, \overline{h}]$ with $\overline{h}$ sufficiently small the points $X^x(t)$ for fixed t and all $x \in \gamma_k$ form a surface $\mathbf{C}^{(3)}$ homeomorphic to γ_k. In this way we obtain an extension of H_k to a neighborhood of U_k by letting $H_k(X^x(t)) = t$ for $t \in [0, \overline{h}]$. The Euclidean distance from a point $X^x(t)$ to γ_k is $\geq \underline{d}\cdot t$ for some $\underline{d} > 0$. Let us denote by $C_k(+t)$ the level surface $\{H_k = +t\}$ for $t \in [0, \overline{h}]$. Let $\underline{\gamma}_k = C_k(+\varepsilon^{1/4})$ and $\underline{\underline{\gamma}}_k = C_k(+2\varepsilon^{1/4})$. Let $\underline{\gamma} = \cup_{k=1}^r \underline{\gamma}_k$ and $\underline{\underline{\gamma}} = \cup_{k=1}^r \underline{\underline{\gamma}}_k$. We can take ε small such that all $\underline{\underline{\gamma}}_k$'s do not intersect each other and do not touch ∂G. We denote by $\mathcal{E}(\varepsilon^{1/4})$ those points of $x \in \mathcal{E}$ which lie outside the union of the neighborhoods of the U_k's bounded by $\underline{\gamma}_k$, and we denote $\mathcal{E}(2\varepsilon^{1/4})$ in a similar way.

We shall denote, for $x \in \mathcal{E}(\varepsilon^{1/4})$, the stopping time $\sigma(\varepsilon^{1/4})$ to be the time when X_t^ε first hits $\underline{\gamma}$. Notice that by our assumption, for a point $x \in \mathcal{E}(\varepsilon^{1/4})$ we have

$$\sum_{i,j=1}^d a_{ij}^{(0)}(x)\xi_i\xi_j \geq \text{const}\cdot\varepsilon^{1/2}\sum_{i,j=1}^d \xi_i^2$$

for all $(\xi_1, ..., \xi_d) \in \mathbb{R}^d$ and some const > 0.

Lemma 4.1. *For any $0 < \varkappa < 1/2$, for any ε small enough we have*

$$\sup_{x\in[\mathcal{E}(\varepsilon^{1/4})]} \mathbf{E}_x\sigma(\varepsilon^{1/4}) \leq C\varepsilon^{1/2-\varkappa}$$

for some $C > 0$.

Proof. Our argument follows from [10, Ch.6]. Let $u^\varepsilon(x,t) = \mathbf{P}_x(\sigma(\varepsilon^{1/4}) > t)$. Then $u^\varepsilon(x,t)$ solves the problem

$$\begin{cases} \dfrac{\partial u^\varepsilon}{\partial t} = \left(\dfrac{1}{\varepsilon}L_0 + L_1\right) u^\varepsilon \ , \\ u^\varepsilon(y,0) = 1 \text{ for } y \in \mathcal{E}(\varepsilon^{1/4}) \ , \\ u^\varepsilon(y,t) = 0 \text{ for } y \in \underline{\gamma} \text{ and } t > 0 \ , \\ \dfrac{\partial u^\varepsilon}{\partial \gamma^\varepsilon}(y,t) = 0 \text{ for } y \in \partial G \ . \end{cases}$$

Let $\varphi(x) = e^{\alpha R} - e^{\alpha x_1}$ for some $\alpha > 0$. Here $R > 0$ is so chosen that $R \geq 2x_1$ for all $x = (x_1, ..., x_d) \in [G]$. We have $\varphi(x) \geq 0$ for $x \in [G]$. We have

$$\begin{aligned} &\left(\frac{1}{\varepsilon}L_0 + L_1 - \frac{\partial}{\partial t}\right)\varphi(x) \\ &= -\frac{1}{2\varepsilon}a_{11}^{(0)}(x)\alpha^2 e^{\alpha x_1} - \frac{1}{2\varepsilon}\frac{\partial a_{11}^{(0)}}{\partial x_1}\alpha e^{\alpha x_1} - \frac{1}{2}a_{11}^{(1)}(x)\alpha^2 e^{\alpha x_1} - \frac{1}{2}\frac{\partial a_{11}^{(1)}}{\partial x_1}\alpha e^{\alpha x_1} \ . \end{aligned}$$

One can choose α large enough but independent of ε such that

$$\left(\frac{1}{\varepsilon}L_0 + L_1 - \frac{\partial}{\partial t}\right)\varphi \leq -\frac{P}{\varepsilon^{1/2}}$$

with $P = \inf\limits_{x\in[G]} e^{\alpha x_1}$.

Let $P_0 = \inf\limits_{x\in[G]} \varphi(x)$ and $P_1 = \sup\limits_{x\in[G]} \varphi(x)$. Consider an auxiliary function

$$\psi(x,t) = \varepsilon\frac{\varphi(x)}{P} + \varepsilon\frac{\varphi(x)}{P_0} + A\frac{\varphi(x)}{P_0}e^{-\beta(t/\varepsilon - \rho)} \ .$$

Here $\rho > 0$ is a small constant (which can be chosen arbitrarily small) such that $u^\varepsilon(y,t) = 0$ for $y \in \underline{\gamma}$ and $t > \rho\varepsilon$. The constants $A > 0$ and $\beta > 0$ are to be chosen later.

We can calculate

$$\begin{aligned} &\left(\frac{1}{\varepsilon}L_0 + L_1 - \frac{\partial}{\partial t}\right)\psi \\ &\leq -\varepsilon^{1/2} - \frac{P}{P_0}\varepsilon^{1/2} - \frac{P}{\varepsilon^{1/2}}\frac{A}{P_0}e^{-\beta(t/\varepsilon-\rho)} + \frac{\beta}{\varepsilon}A\frac{P_1}{P_0}e^{-\beta(t/\varepsilon-\rho)} \ . \end{aligned}$$

Letting $\beta = \beta(\varepsilon) = \dfrac{\varepsilon^{1/2}P}{P_1}$ we have

$$\psi(x,t) = \varepsilon\frac{\varphi(x)}{P} + \varepsilon\frac{\varphi(x)}{P_0} + A\frac{\varphi(x)}{P_0}e^{-\frac{P\varepsilon^{1/2}}{P_1}(t/\varepsilon - \rho)} \ .$$

We have

$$\left(\frac{1}{\varepsilon}L_0 + L_1 - \frac{\partial}{\partial t}\right)\psi \leq -\varepsilon^{1/2} \ .$$

Also $\psi(x,t) > \varepsilon$ for $x \in \underline{\gamma}$ and $t \geq 0$.

Let $A > \sup\limits_{x\in[G]} |u^\varepsilon(x,\rho\varepsilon)|$ so that $\psi(x,\rho\varepsilon) > A > \sup\limits_{x\in[G]} |u^\varepsilon(x,\rho\varepsilon)|$ for $x \in [G]$.

Since

$$\left(\frac{1}{\varepsilon}L_0 + L_1 - \frac{\partial}{\partial t}\right)(\pm u^\varepsilon) = 0$$

and $\pm u^\varepsilon(x,t) = 0 < \varepsilon < \psi(x,t)$ for $x \in \underline{\gamma}$ and $t > 0$, by comparison, we have

$$|u^\varepsilon(x,t)| \leq \psi(x,t) \leq A_1\varepsilon + A_2\exp(-A_3\frac{t}{\varepsilon^{1/2}} + \varepsilon^{1/2}A_4)$$

for some constants $A_1, A_2, A_3, A_4 > 0$ and $x \in [G]$, $t > \rho\varepsilon$.

This implies

$$\sup_{x\in[\mathcal{E}(\varepsilon^{1/4})]} \mathbf{P}_x(\sigma(\varepsilon^{1/4}) > \varepsilon^{1/2-\varkappa}) \leq A_1\varepsilon + A_2\exp(-A_3\varepsilon^{-\varkappa} + \varepsilon^{1/2}A_4) \ .$$

By strong Markov property of the process X_t^ε we see that

$$\begin{aligned}
&\mathbf{E}_x\sigma(\varepsilon^{1/4}) \\
&= \int_0^\infty \mathbf{P}_x(\sigma(\varepsilon^{1/4}) > t)dt \\
&\leq \varepsilon^{1/2-\varkappa}\sum_{n=0}^\infty \mathbf{P}_x(\sigma(\varepsilon^{1/4}) > n\varepsilon^{1/2-\varkappa}) \\
&\leq \varepsilon^{1/2-\varkappa}\sum_{n=0}^\infty (\sup_{x\in[\mathcal{E}(\varepsilon^{1/4})]} \mathbf{P}_x(\sigma(\varepsilon^{1/4}) > \varepsilon^{1/2-\varkappa}))^n \\
&= \frac{\varepsilon^{1/2-\varkappa}}{1 - \sup_{x\in[\mathcal{E}(\varepsilon^{1/4})]} \mathbf{P}_x(\sigma(\varepsilon^{1/4}) > \varepsilon^{1/2-\varkappa})} \\
&\leq \frac{\varepsilon^{1/2-\varkappa}}{1 - A_1\varepsilon - A_2\exp(-A_3\varepsilon^{-\varkappa} + \varepsilon^{1/2}A_4)} \leq C\varepsilon^{1/2-\varkappa}
\end{aligned}$$

for ε small enough. This implies the statement of the Lemma. $\square$

We shall denote by $\mathcal{S}_k([0,\varepsilon^{1/4}])$ the closed set bounded by the surfaces $\underline{\gamma}_k$ and γ_k and by $\mathcal{S}([0,\varepsilon^{1/4}]) = \cup_{k=1}^r \mathcal{S}_k([0,\varepsilon^{1/4}])$. We denote $\mathcal{S}_k([0,2\varepsilon^{1/4}])$ and $\mathcal{S}([0,2\varepsilon^{1/4}])$ in a similar way by replacing $\underline{\gamma}_k$ by $\underline{\underline{\gamma}}_k$.

Following the geometric construction stated before Lemma 4.1, for $\varepsilon > 0$ small enough, and each $k = 1, ..., r$, at any point $x \in \mathcal{S}_k([0,2\varepsilon^{1/4}])$ one can introduce an orthonormal frame $\{\mathbf{e}_j(x)\}_{j=1}^d$ smoothly depending on $x \in \mathcal{S}([0,2\varepsilon^{1/4}])$ such that $\mathbf{e}_d(x) = \dfrac{\nabla H_k(x)}{|\nabla H_k(x)|_{\mathbb{R}^d}}$ and $\mathbf{e}_j(x)\cdot(a^{(0)}(x)\mathbf{e}_j(x)) \geq \underline{a}|\mathbf{e}_j(x)|^2_{\mathbb{R}^d} = \underline{a}$ for some $\underline{a} > 0$ and $j = 1, ..., d-1$. Also $a^{(0)}(x)\mathbf{e}_d(x) = \lambda(x)\mathbf{e}_d(x)$. The eigenvalue $\lambda(x)$ is in $\mathbf{C}^{(3)}(\mathcal{S}_k([0,2\varepsilon^{1/4}]))$ with $\lambda|_{\gamma_k} = 0$ and $\lambda(x) > 0$ for $x \in \mathcal{S}_k([0,2\varepsilon^{1/4}])\backslash\gamma_k$. Furthermore, for ε small enough we have $C_1\cdot\text{dist}^2(x,\gamma_k) \leq \lambda(x) \leq C_2\cdot\text{dist}^2(x,\gamma_k)$ for some $C_1, C_2 > 0$ and $x \in \mathcal{S}([0,2\varepsilon^{1/4}])$.

Within the rest of this section implied positive constants denoted by C_i's will not be explicitly pointed out unless necessary. Also, sometimes we use the same symbol C to denote different implied positive constants which are not important.

Let us introduce a new coordinate $(\varphi_1^k, ..., \varphi_{d-1}^k, H_k)$ in $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$. We take $H_k = H_k(x)$, which is the extended first integral of H_k to $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$; and we take $(\varphi_1^k, ..., \varphi_{d-1}^k) = (\varphi_1^k(x), ..., \varphi_{d-1}^k(x))$ to be the coordinate for a point $\varphi^k(x) = (\varphi_1^k(x), ..., \varphi_{d-1}^k(x))$ on γ_k. The point $\varphi^k(x) \in \gamma_k$ is such that $X^{\varphi^k(x)}(H_k(x)) = x$ for the flow $X^x(t)$ introduced in the geometric construction before Lemma 4.1. In the more or less simpler case we can arrange the coordinate $(\varphi_1^k(x), ..., \varphi_{d-1}^k(x), H_k(x))$ in such a way that $(\mathbf{e}_1(x), ..., \mathbf{e}_d(x))$ is the orthonormal frame corresponding to axis curves of this new coordinate system. (We will discuss the general case a bit later.) The metric tensor of this new coordinate system is given by $ds^2 = E_1(x)(d\varphi_1^k(x))^2 + ... + E_{d-1}(x)(d\varphi_{d-1}^k(x))^2 + E_d(x)(dH_k(x))^2$. Here the functions $0 < C_3 < E_1(x), ..., E_d(x) < C_4 < \infty$ are in class $\mathbf{C}^{(3)}(\mathcal{S}_k([0, 2\varepsilon^{1/4}]))$ with bounded derivatives. We notice that by our geometric construction we have $C_5 \cdot H_k^2(x) \leq \lambda(x) \leq C_6 \cdot H_k^2(x)$.

The theory of orthogonal curvilinear coordinate system (see, for example, [15, Ch.14]) tells us that for a differentiable function f on $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$ we have

$$\nabla f(x) = \sum_{i=1}^{d-1} \frac{1}{\sqrt{E_i(x)}} \frac{\partial f}{\partial \varphi_i^k}(x)\mathbf{e}_i(x) + \frac{1}{\sqrt{E_d(x)}} \frac{\partial f}{\partial H_k}(x)\mathbf{e}_d(x) \ ,$$

and for a differentiable vector field $\mathbf{B}(x) = \sum\limits_{i=1}^{d} B^i(x)\mathbf{e}_i(x)$ on $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$ we have

$$\nabla \cdot \mathbf{B}(x) = \frac{1}{\sqrt{\prod_{i=1}^d E_i(x)}} \left[\sum_{i=1}^{d-1} \frac{\partial}{\partial \varphi_i^k} \left(\sqrt{\frac{\prod_{j=1}^d E_j(x)}{E_i(x)}} B^i(x) \right) + \frac{\partial}{\partial H_k} \left(\sqrt{\frac{\prod_{j=1}^d E_j(x)}{E_d(x)}} B^d(x) \right) \right] .$$

Consider a function (so called "barrier function", see [11] and [4, Ch.3]) $u_k(x) \in \mathbf{C}^{(2)}(\mathcal{S}_k([0, 2\varepsilon^{1/4}]))$ which depends only on H_k and is a constant on each level surface $\{H_k = \text{const}\}$. We can write $u_k(x) = u_k(H_k)$ and we apply the above two formulas to get

$$\begin{aligned}
&\left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k(x) \\
&= \left(\frac{1}{2\varepsilon} \nabla \cdot (a^{(0)}(x) \nabla u_k(x)) + \frac{1}{2} \nabla \cdot (a^{(1)}(x) \nabla u_k(x)) \right) \\
&= \frac{1}{\sqrt{\prod_{i=1}^d E_i(x)}} \left[\frac{1}{2} \frac{\partial}{\partial H_k} \left(\sqrt{\frac{\prod_{i=1}^d E_i(x)}{E_d^2(x)}} \left(\frac{\lambda(x)}{\varepsilon} + \mu_d(x) \right) \frac{du_k}{dH_k}(H_k) \right) \right. \\
&\quad \left. + \sum_{i=1}^{d-1} \frac{1}{2} \frac{\partial}{\partial \varphi_i^k} \left(\sqrt{\frac{\prod_{i=1}^d E_i(x)}{E_d(x) E_i(x)}} \mu_i(x) \right) \cdot \frac{du_k}{dH_k}(H_k) \right] .
\end{aligned}$$

Here the functions $\mu_1(x), ..., \mu_d(x)$ are defined via the relation $a^{(1)}(x)\mathbf{e}_d(x) = \mu_1(x)\mathbf{e}_1(x) + ... + \mu_d(x)\mathbf{e}_d(x)$. These functions are in $\mathbf{C}^{(3)}(\mathcal{S}([0, 2\varepsilon^{1/4}]))$ with bounded derivatives. No-

tice that since L_1 is strictly elliptic, the matrix $a^{(1)}(x)$ is positive definite, and therefore the function $\mu_d(x)$ is uniformly bounded from below by a certain positive constant.

For simplicity of notation let us define $A(x) = \sqrt{\prod_{i=1}^d E_i(x)}$ and $A_i(x) = \dfrac{A(x)}{\sqrt{E_i(x)E_d(x)}}$ for $i = 1, ..., d$. These functions are strictly positive (with uniform lower bound) in $\mathbf{C}^{(2)}(\mathcal{S}([0, 2\varepsilon^{1/4}]))$ with bounded derivatives. Under this notation we can write

$$\left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k(x)$$
$$= \frac{1}{A(x)}\left[\frac{1}{2}\frac{\partial}{\partial H_k}\left(A_d(x)\left(\frac{\lambda(x)}{\varepsilon} + \mu_d(x)\right)\frac{du_k}{dH_k}(H_k)\right) + \sum_{i=1}^{d-1}\frac{1}{2}\frac{\partial}{\partial \varphi_i^k}\left(A_i(x)\mu_i(x)\right)\cdot\frac{du_k}{dH_k}(H_k)\right] .$$

As a further simplification we shall define

$$\frac{1}{2}A_d(x)\lambda(x) = K_1(x) ,$$

$$\frac{1}{2}A_d(x)\mu_d(x) = K_2(x) ,$$

$$\sum_{i=1}^{d-1}\frac{1}{2}\frac{\partial}{\partial \varphi_i^k}(A_i(x)\mu_i(x)) = K_3(x) .$$

We have

$$\left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k(x) = \frac{1}{A(x)}\left[\frac{\partial}{\partial H_k}\left(\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right)\frac{du_k}{dH_k}(H_k)\right) + K_3(x)\frac{du_k}{dH_k}(H_k)\right] . \tag{4.1}$$

For a point $x \in \mathcal{S}_k([0, 2\varepsilon^{1/4}])$ and ε small enough we have

$$C_7 H_k^2(x) \leq K_1(x) \leq C_8 H_k^2(x) ; \tag{4.2}$$

$$C_9 H_k(x) \leq \frac{\partial}{\partial H_k}(K_1(x)) \leq C_{10} H_k(x) ; \tag{4.3}$$

$$0 < C_{11} < K_2(x) < C_{12} < \infty ; \tag{4.4}$$

$$\left|\frac{\partial}{\partial H_k}(K_2(x))\right| \leq C_{13} < \infty ; \tag{4.5}$$

$$|K_3(x)| \leq C_{14} < \infty . \tag{4.6}$$

We also notice, that since we are working in a small neighborhood $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$, the functions $A_d(x) = A_d(\varphi_1^k, ..., \varphi_{d-1}^k, H_k)$ and $\lambda(x) = \lambda(\varphi_1^k, ..., \varphi_{d-1}^k, H_k)$ have Taylor expansions

$$A_d(\varphi_1^k, ..., \varphi_{d-1}^k, H_k) = A_d(\varphi_1^k, ..., \varphi_{d-1}^k, 0) + O(H_k) ,$$

$$\lambda(\varphi_1^k, ..., \varphi_{d-1}^k, H_k) = \frac{1}{2}\frac{\partial^2 \lambda}{\partial H_k^2}(\varphi_1^k, ..., \varphi_{d-1}^k, 0)H_k^2 + O(H_k^3) .$$

Therefore we see that for $x \in \mathcal{S}_k([0, 2\varepsilon^{1/4}])$ we have

$$K_1(x) = C_k(\varphi_1^k, ..., \varphi_{d-1}^k)H_k^2 + O(H_k^3) \tag{4.7}$$

with a certain positive function $C_k(\varphi_1^k, ..., \varphi_{d-1}^k)$.

In the general case the axis curve corresponding to H_k will be orthogonal to those corresponding to the φ_i^k's, but the axis curves corresponding to the φ_i^k's are not necessarily orthogonal. The calculation will be more bulky since the metric tensor have cross terms with respect to the coordinate φ_i^k's, but the essence is the same as it is only important to have the axis curves corresponding to H_k being orthogonal to those corresponding to the φ_i^k's. To be more precise, let $(g_{ij})_{1\leq i,j\leq d}$ be the metric tensor corresponding to the coordinate system $(\varphi_1^k, ..., \varphi_{d-1}^k, H_k)$. We introduce a frame $\mathbf{e}_1(x), ..., \mathbf{e}_d(x)$. Here $\mathbf{e}_i(x)$ is the unit tangent vector on the axis curve corresponding to φ_i^k for $1 \leq i \leq d-1$; $\mathbf{e}_d(x)$ is the unit tangent vector on the axis curve corresponding to H_k. We have $g_{id} = g_{di} = 0$ for $i = 1, ..., d-1$ and $g_{dd} > 0$. Let $(g^{ij})_{1\leq i,j\leq d}$ be the dual tensor, i.e., $(g^{ij})_{1\leq i,j\leq d}$ is the inverse matrix of $(g_{ij})_{1\leq i,j\leq d}$. We have $g^{id} = g^{di} = 0$ for $1 \leq i \leq d-1$ and $g^{dd} = \dfrac{1}{g_{dd}}$. For $u_k = u_k(H_k)$ we have

$$\nabla u_k(x) = \frac{1}{\sqrt{g_{dd}(x)}}\frac{du_k}{dH_k}\mathbf{e}_d(x) \ ,$$

and

$$a^{(0)}(x)\nabla u_k(x) = \frac{\lambda(x)}{\sqrt{g_{dd}(x)}}\frac{du_k}{dH_k}\mathbf{e}_d(x) \ ;$$

$$a^{(1)}(x)\nabla u_k(x) = \frac{\mu_d(x)}{\sqrt{g_{dd}(x)}}\frac{du_k}{dH_k}\mathbf{e}_d(x) + \frac{1}{\sqrt{g_{dd}(x)}}\frac{du_k}{dH_k}(\mu_1(x)\mathbf{e}_1(x) + ... + \mu_{d-1}(x)\mathbf{e}_{d-1}(x)) \ .$$

Here, as before, we have $a^{(1)}(x)\mathbf{e}_d(x) = \mu_1(x)\mathbf{e}_1(x) + ... + \mu_d(x)\mathbf{e}_d(x)$. We shall then apply a general formula that for a vector field $\mathbf{B}(x) = \sum\limits_{i=1}^{d} B^i(x)\mathbf{e}^i(x)$ we have

$$\nabla \cdot \mathbf{B}(x) = \frac{1}{\sqrt{g(x)}}\sum_{i=1}^{d-1}\frac{\partial}{\partial \varphi_i^k}(B^i(x)\sqrt{g^{ii}(x)}\sqrt{g(x)}) + \frac{1}{\sqrt{g(x)}}\frac{\partial}{\partial H_k}(B^d(x)\sqrt{g^{dd}(x)}\sqrt{g(x)}) \ .$$

Here $g(x) = \det(g_{ij}(x))$. The basis $\mathbf{e}^1(x), ..., \mathbf{e}^d(x)$ is the reciprocal basis (normalized) dual to $\mathbf{e}_1(x), ..., \mathbf{e}_d(x)$, i.e., $(\mathbf{e}_i, \mathbf{e}^j)_{(g_{ij})} = \delta_{ij}$ with respect to the inner product $(\bullet, \bullet)_{(g_{ij})}$ defined by the metric tensor (g_{ij}). By the fact that the metric tensor has no cross terms between H_k and φ_i^k's, we actually have $\mathbf{e}^d(x) = \mathbf{e}_d(x)$ and $\text{span}\{\mathbf{e}_1(x), ..., \mathbf{e}_{d-1}(x)\} = \text{span}\{\mathbf{e}^1(x), ..., \mathbf{e}^{d-1}(x)\}$.

We then see that the operator $\dfrac{1}{\varepsilon}L_0 + L_1$ applied to $u_k(x) = u_k(H_k)$ will result in a formula which is the same as (4.1). The functions $K_1(x)$, $K_2(x)$ and $K_3(x)$ will somehow be different but they still satisfy the conditions (4.2) − (4.7).

Let $\zeta([0,2\varepsilon^{1/4}])$ be the first time when the process X_t^ε, starting from a point $x \in \mathcal{S}([0,2\varepsilon^{1/4}])$, hits γ or $\underline{\underline{\gamma}}$.

Lemma 4.2. *We have*

$$\sup_{x\in\mathcal{S}([0,2\varepsilon^{1/4}])} \mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) \leq \ C\varepsilon^{3/4}$$

for some $C > 0$.

Proof. Let

$$K_4(H_k) = \min_{x\in\mathcal{S}_k([0,2\varepsilon^{1/4}]),H_k(x)=H_k} \left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right) \ .$$

By (4.2) and (4.4) we can estimate

$$C_{15}\left(\frac{H_k^2}{\varepsilon}+1\right) \leq K_4(H_k) \leq C_{16}\left(\frac{H_k^2}{\varepsilon}+1\right) \tag{4.8}$$

for $C_{15}, C_{16} > 0$.

Let

$$K_5(x) = \frac{1}{K_4(H_k(x))}\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right) \geq 1 \ .$$

The function $K_5(x)$ is a bounded function with bounded derivatives for $x \in \mathcal{S}([0,2\varepsilon^{1/4}])$.

Let the barrier function $u_k^{(1)}(x) = u_k^{(1)}(H_k)$ be defined by

$$u_k^{(1)}(H_k) = \int_0^{H_k} \frac{K_6(\varepsilon) - y}{K_4(y)} dy$$

with

$$K_6(\varepsilon) = \left(\int_0^{2\varepsilon^{1/4}} \frac{dy}{K_4(y)}\right)^{-1} \left(\int_0^{2\varepsilon^{1/4}} \frac{ydy}{K_4(y)}\right) \ .$$

It is easy to check that

$$u_k^{(1)}(0) = u_k^{(1)}(2\varepsilon^{1/4}) = 0 \ .$$

We can estimate $K_6(\varepsilon) \leq 2\varepsilon^{1/4}$ and we have, by (4.8), that

$$\int_0^{H_k} \frac{dy}{K_4(y)} \leq C_{17}\int_0^{H_k} \frac{dy}{\dfrac{y^2}{\varepsilon}+1} \leq C_{17}\varepsilon^{1/2}\arctan(H_k\varepsilon^{-1/2}) \leq C_{18}\varepsilon^{1/2} \ . \tag{4.9}$$

This gives the estimates

$$0 \leq u_k^{(1)}(H_k) \leq C_{19}\varepsilon^{3/4} \tag{4.10}$$

and

$$\left|\frac{du_k^{(1)}}{dH_k}(H_k)\right| \leq C_{20}\varepsilon^{1/4} \tag{4.11}$$

for $0 \leq H_k \leq 2\varepsilon^{1/4}$. Apply (4.1) to the function $u_k^{(1)}$ we can see, using (4.11), that,

$$\begin{aligned}
&\left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k^{(1)}(x) \\
&= \frac{1}{A(x)}\left[\frac{\partial}{\partial H_k}\left(K_5(x)K_4(H_k(x))\frac{du_k^{(1)}}{dH_k}(H_k(x))\right) + K_3(x)\frac{du_k^{(1)}}{dH_k}(H_k(x))\right] \\
&\leq \frac{1}{A(x)}\left[\frac{\partial}{\partial H_k}\left((K_6(\varepsilon) - H_k(x))K_5(x)\right) + C_{21}\varepsilon^{1/4}\right] \\
&= \frac{1}{A(x)}\left[-K_5(x) + \frac{\partial}{\partial H_k}(K_5(x))(K_6(\varepsilon) - H_k(x)) + C_{21}\varepsilon^{1/4}\right] \leq -C_{22}
\end{aligned} \tag{4.12}$$

for $x \in \mathcal{S}([0, 2\varepsilon^{1/4}])$ and ε small enough.

We notice that this process X_t^ε before hitting γ or $\underline{\underline{\gamma}}$ is restricted to one of the $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$'s and the bound (4.12) can be made uniform in k.

Now we apply Itô's formula to the function $u_k^{(1)}$ constructed above up to the stopping time $\zeta([0, 2\varepsilon^{1/4}])$. Taking expectation we get

$$u_k^{(1)}(x) = -\int_0^{\zeta([0,2\varepsilon^{1/4}])} \mathbf{E}_x \left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k^{(1)}(X_s^\varepsilon)ds \geq C_{22}\mathbf{E}_x\zeta([0, 2\varepsilon^{1/4}]) \ . \tag{4.13}$$

From (4.10) and (4.13) we see that the statement of this Lemma follows. □

Lemma 4.3. *For $x \in \underline{\gamma}$ we have*

$$\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \gamma) \geq C\varepsilon^{1/4}$$

for some $C > 0$.

Proof. Let

$$K_7(H_k) = \max_{x\in\mathcal{S}_k([0,2\varepsilon^{1/4}]), H_k(x)=H_k} \frac{\dfrac{\partial}{\partial H_k}\left(\dfrac{K_1(x)}{\varepsilon} + K_2(x)\right) + K_3(x)}{\dfrac{K_1(x)}{\varepsilon} + K_2(x)} \ .$$

Let, for a fixed $H_k \in [0, 2\varepsilon^{1/4}]$, the above maximum be achieved at a point

$$\varphi^k = (\varphi_1^k(H_k), ..., \varphi_{d-1}^k(H_k), H_k) \ .$$

By Lemma 4.4 we have

$$\left| K_7(H_k) - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \right| \leq C_{23} \ . \tag{4.14}$$

Let the barrier function $u_k^{(2)}(x) = u_k^{(2)}(H_k)$ be defined by

$$u_k^{(2)}(H_k) = 1 - \frac{\displaystyle\int_0^{H_k} \exp\left(-\int_0^y K_7(z)dz\right) dy}{\displaystyle\int_0^{2\varepsilon^{1/4}} \exp\left(-\int_0^y K_7(z)dz\right) dy} \ .$$

It is easy to see that we have

$$u_k^{(2)}(0) = 1 \ , \ u_k^{(2)}(2\varepsilon^{1/4}) = 0 \ .$$

Apply formula (4.1) we can see that

$$\begin{aligned} &\left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k^{(2)}(x) \\ &= \frac{1}{A(x)} \left[\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right) \frac{d^2 u_k^{(2)}}{dH_k^2}(H_k) + \left(\frac{\partial}{\partial H_k}\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right) + K_3(x)\right) \frac{du_k^{(2)}}{dH_k}(H_k)\right] \\ &\geq 0 \ . \end{aligned} \tag{4.15}$$

However, by (4.14), (4.4) and the property of $C_k(\varphi^k)$ in (4.7) we can estimate

$$\int_{H_k}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y K_7(z)dz\right) dy \geq C_{24} \int_{H_k}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C_{25}\varepsilon} dz\right) dy \ , \tag{4.16}$$

$$\int_0^{2\varepsilon^{1/4}} \exp\left(-\int_0^y K_7(z)dz\right) dy \leq C_{26} \int_0^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C_{27}\varepsilon} dz\right) dy \ . \tag{4.17}$$

By Lemma 4.5, (4.16) and (4.17) we see that

$$u_k^{(2)}(\varepsilon^{1/4}) \geq C_{28}\varepsilon^{1/4} \ . \tag{4.18}$$

This bound (4.18) can actually be made uniform in k. We can apply Itô's formula to the function $u_k^{(2)}$ constructed above up to the stopping time $\zeta([0, 2\varepsilon^{1/4}])$. Taking expectation we get

$$\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \gamma) - u_k^{(2)}(\varepsilon^{1/4}) = \int_0^{\zeta([0,2\varepsilon^{1/4}])} \mathbf{E}_x \left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k^{(2)}(X_s^\varepsilon) ds \geq 0 \tag{4.19}$$

for $x \in \underline{\gamma}$. Now (4.18) and (4.19) imply the statement of this Lemma. $\square$

Lemma 4.4. *For a fixed $H_k \in [0, 2\varepsilon^{1/4}]$ and the corresponding φ^k defined as in the proof of Lemma 4.3, we have*

$$\left| K_7(H_k) - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \right| \le C$$

for some $C > 0$.

Proof. Using (4.7), we can write

$$K_7(H_k) = \frac{2C_k(\varphi^k)H_k + O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} \ .$$

Here $K_{2,3}(\varphi^k, H_k)$ is a bounded function. We then have

$$\begin{aligned}
&\left| K_7(H_k) - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \right| \\
&= \left| \frac{2C_k(\varphi^k)H_k + O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \right| \\
&\le \left| \frac{O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} \right| + \\
&\quad + \left| \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \right| \\
&= \left| \frac{O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} \right| + \\
&\quad + \left| \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \cdot \frac{O(H_k^3)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} \right| \le C \ .
\end{aligned}$$

□

Lemma 4.5. *We have*

$$\int_{\varepsilon^{1/4}}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon} dz \right) dy \ge C_{29}\varepsilon^{3/4} \ ,$$

$$C_{31}\varepsilon^{1/2} \ge \int_0^{C_{32}\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon} dz \right) dy \ge C_{30}\varepsilon^{1/2} \ .$$

Proof. Evaluating the integrals, we have

$$\int_0^y \frac{2z}{z^2 + C\varepsilon} dz = \ln\left(\frac{y^2 + C\varepsilon}{C\varepsilon} \right) \ ,$$

$$\int_a^b \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon} dz \right) dy = \sqrt{C}\varepsilon^{1/2} \left(\arctan(\frac{b}{\sqrt{C}\varepsilon^{1/2}}) - \arctan(\frac{a}{\sqrt{C}\varepsilon^{1/2}}) \right) \ .$$

If $a = 0$ and $b = C_{32}\varepsilon^{1/4}$ we already get the second inequality of this Lemma. Now suppose $a = \varepsilon^{1/4}$ and $b = 2\varepsilon^{1/4}$. We shall make use of an asymptotic expansion of $\arctan(y)$ as $y \to \infty$:

$$\arctan(y) = \frac{\pi}{2} - \frac{1}{y} + O(\frac{1}{y^2}) \quad \text{as} \quad y \to \infty \ .$$

This gives

$$\begin{aligned}
&\int_{\varepsilon^{1/4}}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon} dz\right) dy \\
&\geq C_{33}\varepsilon^{1/2}\left(-\frac{1}{2\varepsilon^{-1/4}} + \frac{1}{\varepsilon^{-1/4}} + O(\varepsilon^{1/2})\right) \geq C_{29}\varepsilon^{3/4} \ .
\end{aligned}$$

□

Lemma 4.6. *We have*

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \underline{\gamma}} \mathbf{E}_x \sigma = 0$$

uniformly in ε.

Proof. Lemmas 4.1, 4.2 and 4.3 imply the statement of this Lemma. For $x \in \underline{\gamma}$ we have

$$\begin{aligned}
&\mathbf{E}_x\sigma \\
&= \mathbf{E}_x\zeta([0,2\varepsilon^{1/4}])\mathbf{1}(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \gamma) + \mathbf{E}_x(\zeta([0,2\varepsilon^{1/4}]) + \mathbf{E}_{X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}}\sigma)\mathbf{1}(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \underline{\underline{\gamma}}) \\
&\leq \sup_{x \in \mathcal{S}([0,2\varepsilon^{1/4}])} \mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) + \mathbf{E}_x(\sup_{y \in \underline{\underline{\gamma}}} \mathbf{E}_y\sigma(\varepsilon^{1/4}) + \sup_{x \in \underline{\gamma}} \mathbf{E}_x\sigma)\mathbf{1}(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \underline{\underline{\gamma}}) \\
&\leq \sup_{x \in \mathcal{S}([0,2\varepsilon^{1/4}])} \mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) + (\sup_{y \in \underline{\underline{\gamma}}} \mathbf{E}_y\sigma(\varepsilon^{1/4}) + \sup_{x \in \underline{\gamma}} \mathbf{E}_x\sigma)\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \underline{\underline{\gamma}}) \ .
\end{aligned} \tag{4.20}$$

Taking a sup over all $x \in \underline{\gamma}$ we get

$$\sup_{x \in \underline{\gamma}} \mathbf{E}_x\sigma \leq \frac{\sup_{x \in \mathcal{S}([0,2\varepsilon^{1/4}])} \mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) + \sup_{y \in \underline{\underline{\gamma}}} \mathbf{E}_y\sigma(\varepsilon^{1/4}) \cdot \mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \underline{\underline{\gamma}})}{\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \gamma)} \ .$$

Using Lemmas 4.1, 4.2 and 4.3 we see that the statement of this Lemma follows. (We choose $\varkappa = 1/8$ in Lemma 4.1.) □

Lemma 4.7. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathcal{S}([0,\varepsilon^{1/4}])}\mathbf{E}_x\sigma=0$$

uniformly in ε.

Proof. This is a consequence of Lemmas 4.1, 4.2 and 4.6. $\square$

Lemma 4.8. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in[\mathcal{E}]}\mathbf{E}_x\sigma=0$$

uniformly in ε.

Proof. This is a consequence of Lemmas 4.1 and 4.7. $\square$

We shall denote by $\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$ the closed set bounded by the surfaces $\overline{\gamma}_k$ and $\underline{\gamma}_k$ and by $\mathcal{S}([-\varepsilon^{1/2},\varepsilon^{1/4}])=\cup_{k=1}^r\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$. We notice that by the same reason as before, a coordinate $(\varphi_1^k,...,\varphi_{d-1}^k,H_k)$ exists in $\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$. We denote $\mathcal{S}_k([0,\varepsilon^{1/4}])$, $\mathcal{S}([0,\varepsilon^{1/4}])$ (replacing $\overline{\gamma}_k$ by γ_k) and $\mathcal{S}_k([-\varepsilon^{1/2},0])$, $\mathcal{S}([-\varepsilon^{1/2},0])$ (replacing $\underline{\gamma}_k$ by γ_k) in a similar way.

Lemma 4.9. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\gamma}\mathbf{E}_x\tau=0$$

uniformly in ε.

Proof. The proof of this lemma is very similar to and is a bit simpler than that of Lemma 4.6. We shall construct two barrier functions $u_k^{(3)}$ (for estimating the exit time from $\mathcal{S}_k([-\varepsilon^{-1/2},\varepsilon^{1/4}]))$ and $u_k^{(4)}$ (for the probability of hitting $\overline{\gamma}$).

For the construction of $u_k^{(3)}$ all the arguments of Lemma 4.2 can be carried here with γ replaced by $\overline{\gamma}$, $\underline{\gamma}$ replaced by γ and $\underline{\underline{\gamma}}$ replaced by $\underline{\gamma}$. We are working now with $\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$ and $H_k\in[-\varepsilon^{1/2},\varepsilon^{1/4}]$. We apply formula (4.1) with the change of the estimates (4.2) – (4.6) as follows: when $x\in\mathcal{S}_k([0,\varepsilon^{1/4}])$ there is no change in the estimates; when $x\in\mathcal{S}_k([-\varepsilon^{1/2},0])$ we replace (4.2) and (4.3) by $K_1(x)=\dfrac{\partial}{\partial H_k}(K_1(x))=0$ and (4.4) – (4.6) remain the same. The function $K_4(x)$ is then defined in a same way as in Lemma 4.2 with an estimate $0<C_{34}\le K_4(H_k)\le C_{35}<\infty$ for $x\in\mathcal{S}_k([-\varepsilon^{1/2},0])$. Again we let

$$u_k^{(3)}(H_k)=\int_{-\varepsilon^{1/2}}^{H_k}\frac{K_6(\varepsilon)-y}{K_4(y)}dy$$

with

$$K_6(\varepsilon) = \left(\int_{-\varepsilon^{1/2}}^{\varepsilon^{1/4}} \frac{dy}{K_4(y)}\right)^{-1} \left(\int_{-\varepsilon^{1/2}}^{\varepsilon^{1/4}} \frac{ydy}{K_4(y)}\right) .$$

It is then checked that $u_k(-\varepsilon^{1/2}) = u_k(\varepsilon^{1/4}) = 0$ and $K_6(\varepsilon) \leq 2\varepsilon^{1/4}$. The estimate (4.9) is still working for $H_k \in [-\varepsilon^{1/2}, \varepsilon^{1/4}]$. The estimates (4.10), (4.11) and (4.12) are still working. Let $\zeta([-\varepsilon^{1/2}, \varepsilon^{1/4}])$ be the first time when the process X_t^ε starting from a point $x \in \gamma$, first hits $\underline{\gamma}$ or $\overline{\gamma}$. A similar statement of (4.13) is then obtained. We have

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \mathcal{S}([-\varepsilon^{1/2}, \varepsilon^{1/4}])} \mathbf{E}_x \zeta([-\varepsilon^{1/2}, \varepsilon^{1/4}]) = 0 \tag{4.21}$$

uniformly in ε.

The estimate of the hitting probability is a bit simpler. We construct a barrier function $u_k^{(4)}$ similarly as in Lemma 4.3. The function $K_7(H_k)$ is defined as in Lemma 4.3. But now we have the property that $|K_7(H_k)| \leq C$ for $H_k \in [-\varepsilon^{1/2}, 0]$. We let

$$u_k^{(4)}(H_k) = 1 - \frac{\displaystyle\int_{-\varepsilon^{1/2}}^{H_k} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy}{\displaystyle\int_{-\varepsilon^{1/2}}^{\varepsilon^{1/4}} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy} .$$

We have $u_k^{(4)}(-\varepsilon^{1/2}) = 1$, $u_k^{(4)}(\varepsilon^{1/4}) = 0$. We have

$$C_{36}\varepsilon^{1/2} \leq \int_{-\varepsilon^{1/2}}^{0} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy \leq C_{37}\varepsilon^{1/2} ,$$

and by the second inequality in Lemma 4.5 we see that

$$C_{38}\varepsilon^{1/2} \leq \int_{0}^{\varepsilon^{1/4}} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy \leq C_{39}\varepsilon^{1/2} .$$

These estimates ensure that an analogue of (4.19) works, but the lower bound is a positive constant, and hence situation is a bit simpler. We have

$$1 \geq \mathbf{P}_x(X^\varepsilon_{\zeta([-\varepsilon^{1/2}, \varepsilon^{1/4}])} \in \overline{\gamma}) \geq u_k^{(4)}(0) \geq C_{40} > 0 . \tag{4.22}$$

uniformly in $x \in \gamma$ and $\varepsilon > 0$. The results (4.21), (4.22) and Lemma 4.8, combined with a similar analysis of (4.20) in Lemma 4.6, give the statement of this Lemma. $\square$

Lemma 4.10. *We have*

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \overline{\gamma}} \mathbf{E}_x \sigma = 0$$

uniformly in ε.

Proof. This is a result in the same essence of Lemma 3.2 (formula (10)) of [2]. □

Lemma 4.11. *The process $X^{\varepsilon}_{\sigma_n}$ satisfies the Doeblin condition on γ uniformly in ε.*

Proof. For each fixed $\varepsilon > 0$ we have the ergodicity of the process X^{ε}_t. Uniformly in ε the Doeblin condition is satisfied for the process X^{ε}_t in $[\mathcal{E}]$ and each of these $[U_k]$'s for $k = 1, ..., r$. As we have Lemmas 4.8, 4.9 and 4.10, we see that the statement of this Lemma follows. □

Acknowledgement. The second author would like to thank Konstantinos Spiliopoulos for reading the first version of this paper and for providing many valuable suggestions on the improvement of the presentation.

References

[1] Dolgopyat, D., Koralov, L., Averaging of Hamiltonian flows with an ergodic component, *Annals of Probability*, **36** (2008), pp. 1999 – 2049.

[2] Dolgopyat, D., Koralov, L., Averaging of incompressible flows on two dimensional surfaces, preprint, submitted to *Journal of AMS*.

[3] Freidlin, M., On the factoriztion of a non-negative definite matrix, *Theory of Probability and its Applications*, **13** (1968), pp. 354 – 356 (English translation).

[4] Freidlin, M., *Functional integration and partial differential equations*, Princeton University Press, 1985.

[5] Freidlin, M., Weber, M., Random perturbations of dynamical systems and diffusion processes with conservation laws, *Probability Theory and Related Fields*, **128**, pp. 441 – 466, 2004.

[6] Freidlin, M., Wentzell, A., *Random perturbations of dynamical systems*, Second Edition, Springer, 1998.

[7] Freidlin, M., Wentzell, A., Diffusion processes on graphs and the averaging principle, *Annals of Probability*, **21**, 4, 1993, pp. 2215 – 2245.

[8] Freidlin, M., Wentzell, A., Random Perturbations of Hamiltonian Systems, *Mem. of AMS*, **109** (1994).

[9] Freidlin, M., Wentzell, A., On the Neumann problem for PDE's with a small parameter and corresponding diffusion processes, *Probability Theory and Related Fields*, **152**, pp. 101 − 140, 2012.

[10] Freidman, A., *Partial differential equations of parabolic type*, Englewood Cliffs, N.J., Prentice-Hall, 1964.

[11] Has'minskii, R.Z., Diffusion processes and elliptic differential equations degenerating on the boundary of the region, *Theory of Probability and its Applications*, **3** (1958), pp. 430 − 461 (English translation).

[12] Itô, K., McKean, H.P. Jr., Brownian motion on a half line, *Illinois Journal of Mathematics*, **7**, pp. 181 − 231, 1963.

[13] Itô, K., McKean, H.P. Jr., *Diffusion processes and their sample paths*, Second Edition, Springer, 1974.

[14] Levinson, N., The first boundary-value problem for $\varepsilon\Delta u + Au_x + Bu_y + Cu = 0$ for small ε, *Annals of Mathematics*, **51**, pp. 428 − 445, 1950.

[15] Zorich, V.A., *Mathematical Analysis*, Part II, 4th corrected edition, Moscow, MCCME, 2002 (in Russian), Chinese translation, Higher Education Press, Beijing, 2006.